

205

Arg Lys Val Leu Glu Leu Ala Ala Ala Leu Ser Asp Asp Phe Glu Arg
290 295 300

Ala Gly Arg Arg
305

(2) INFORMATION FOR SEQ ID NO: 95:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 95:

AAGACTAGAT CTATGATGGC CGAGGATGTT CGCG

24

(2) INFORMATION FOR SEQ ID NO: 96:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 96:

CGCGGACGAC GGATCTTACC GCGTGGG

27

(2) INFORMATION FOR SEQ ID NO: 97:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 97:

CCTTGGGAGA TCTTTGGACC CCGGTTC

28

(2) INFORMATION FOR SEQ ID NO: 98:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

{xi} SEQUENCE DESCRIPTION: SEQ ID NO: 98:

GACGAGATCT TATGGGCTTA CTGAC

25

{2} INFORMATION FOR SEQ ID NO: 99:

{1} SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

{xi} SEQUENCE DESCRIPTION: SEQ ID NO: 99:

CCCCCAGAT CTGCACCACC GGCATCGGCG GGC

33

{2} INFORMATION FOR SEQ ID NO: 100:

{1} SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

{xi} SEQUENCE DESCRIPTION: SEQ ID NO: 100:

GCGGCGGATC CGTTGCTTAG CCGG

24

{2} INFORMATION FOR SEQ ID NO: 101:

{1} SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

{xi} SEQUENCE DESCRIPTION: SEQ ID NO: 101:

CCGGCTGAGA TCTATGACAG AATACGAGG GC

32

{2} INFORMATION FOR SEQ ID NO: 102:

{1} SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

{xi} SEQUENCE DESCRIPTION: SEQ ID NO: 102:

CCCCCCCCG GAAC TAGAGG CCGC

24

{2} INFORMATION FOR SEQ ID NO: 103:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 103:

CTGCCGAGAT CTACCACCAT TGTGCGGCTG AAATACCC

38

(2) INFORMATION FOR SEQ ID NO: 104:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 104:

CGCCATGGCC TTACGCGCCA ACTCG

25

(2) INFORMATION FOR SEQ ID NO: 105:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 105:

GGCGGAGATC TGTGAGTTTT CCGTATTTC TC

32

(2) INFORMATION FOR SEQ ID NO: 106:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 106:

CGCGTCGAGC CATGGTTAGG CGCAG

25

(2) INFORMATION FOR SEQ ID NO: 107:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 107:

GAGGAAGATC TATGACAACT TCACCCGACC CG 32

(2) INFORMATION FOR SEQ ID NO: 108:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 108:

CATGAAGCCA TGGCCCGCAG GCTGCATG 28

(2) INFORMATION FOR SEQ ID NO: 109:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 109:

GGCCGAGATC TGTGACCCAC TATGACGTCC TCG 33

(2) INFORMATION FOR SEQ ID NO: 110:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 110:

GGCGCCCATG GTCAGAAAT GATCATGTGG CCAACC 36

(2) INFORMATION FOR SEQ ID NO: 111:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 111:

CCGGGAGATC TATGGCAAAG CTCGCCACCG ACC 33

(2) INFORMATION FOR SEQ ID NO: 112:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 112:

CGCTGGGCAG AGCTACTTGA CGGTGACGGT GG 32

(2) INFORMATION FOR SEQ ID NO: 113:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 113:

GGCCGAGATC TATGGCCATT GAGTTTCGG TGTGGC 36

(2) INFORMATION FOR SEQ ID NO: 114:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 114:

CGCCGTGTTC CATGGCAGCG CTGAGC 26

(2) INFORMATION FOR SEQ ID NO: 115:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 115:

GGACGTTCAA GCGACACATC GCGC 24

(2) INFORMATION FOR SEQ ID NO: 116:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 116:

CAGCAGGAAC GCGCCGTGCA TGGC

24

(2) INFORMATION FOR SEQ ID NO: 117:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 117:

ACAGATCTGT GACGGACATG AACCCG

25

(2) INFORMATION FOR SEQ ID NO: 118:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 118:

TTTCCATGG TCACGGGCCC CCGTACT

26

(2) INFORMATION FOR SEQ ID NO: 119:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 119:

ACAGATCTGT GCCCATGGCA CAGATA

26

(2) INFORMATION FOR SEQ ID NO: 120:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 120:

TTTAAGCTTC TAGGCGCCCA GCGCGGC

27

(2) INFORMATION FOR SEQ ID NO: 121:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 121:

ACAGATCTGC GCATGCGGAT CCGTGT 26

(2) INFORMATION FOR SEQ ID NO: 122:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 38 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 122:

TTTCCATGG TCAICCGGGG TGATCGAG 28

(2) INFORMATION FOR SEQ ID NO: 123:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 123:

ACAGATCTGT AATGGCAGAC TGTGAT 26

(2) INFORMATION FOR SEQ ID NO: 124:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 124:

TTTCCATGG TCAGGAGATG GTGATCGA 28

(2) INFORMATION FOR SEQ ID NO: 125:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

212

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 125:

ACAGATCTGC CGGCTACCCC GGTGCC

26

(2) INFORMATION FOR SEQ ID NO: 126:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 126:

TTTCCATGG CTATTGCAGC TTCCGGC

28

(2) INFORMATION FOR SEQ ID NO: 127:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 50 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 127:

Ala	Glu	Asp	Val	Arg	Ala	Glu	Ile	Val	Ala	Ser	Val	Leu	Glu	Val	Val
1				5					10					15	
Val	Asn	Glu	Gly	Asp	Gln	Ile	Asp	Lys	Gly	Asp	Val	Val	Val	Leu	Leu
			20					25						30	
Glu	Ser	Met	Tyr	Met	Glu	Ile	Pro	Val	Leu	Ala	Glu	Ala	Ala	Gly	Thr
			35					40						45	
Val	Ser														
			50												

(2) INFORMATION FOR SEQ ID NO: 128:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 49 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 128:

Ala	Glu	Asp	Val	Arg	Ala	Glu	Ile	Val	Ala	Ser	Val	Leu	Glu	Val	Val
1				5					10					15	
Val	Asn	Glu	Gly	Asp	Gln	Ile	Asp	Lys	Gly	Asp	Val	Val	Val	Leu	Leu
			20					25						30	

213

Glu Ser Met Met Glu Ile Pro Val Leu Ala Glu Ala Ala Gly Thr Val
 35 40 45

Ser

(2) INFORMATION FOR SEQ ID NO: 129:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 50 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 129:

Ala Glu Asp Val Arg Ala Glu Ile Val Ala Ser Val Leu Glu Val Val
 1 5 10 15

Val Asn Glu Gly Asp Gln Ile Asp Lys Gly Asp Val Val Val Leu Leu
 20 25 30

Glu Ser Met Lys Met Glu Ile Pro Val Leu Ala Glu Ala Ala Gly Thr
 35 40 45

Val Ser
 50

(2) INFORMATION FOR SEQ ID NO: 130:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 130:

CCGGGAGATC TATGGCAAAG CTCTCCACCG ACC

33

(2) INFORMATION FOR SEQ ID NO: 131:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 131:

CCCTGGGCGAG AGCTACTTGA CGGTGACGGT GG

32

(2) INFORMATION FOR SEQ ID NO: 132:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 36 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 132:

GGGCCCCGGCA AGCTTGCCAT GACAGAGCAG CAGTGG 36

(2) INFORMATION FOR SEQ ID NO: 133:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 26 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 133:

CGAACTGGCC GGATCCCGTG TTTCGC 26

(2) INFORMATION FOR SEQ ID NO: 134:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 32 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 134:

GGCRAUCCGCG AGATCTTTCT CCGGSCCGGG GC 32

(2) INFORMATION FOR SEQ ID NO: 135:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 27 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 135:

GGCRACTTG CCGGCGCCTA ACGAACT 27

(2) INFORMATION FOR SEQ ID NO: 136:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 30 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

215

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 136:

GGACCCAGAT CTATGACAGA GCAGCAGTGG

30

(2) INFORMATION FOR SEQ ID NO: 137:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 47 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 137:

CCGGCAGCCC CGGCCGGGAG AAAAGCTTTG CGACATCCC AGTGACC

47

(2) INFORMATION FOR SEQ ID NO: 138:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 44 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 138:

GTTCCGAAG CTTTCTCCC GGCCGGGCT GCGGTCCAG TACC

44

(2) INFORMATION FOR SEQ ID NO: 139:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 139:

CCTTCGGTGG ATCCGTCAG

20

(2) INFORMATION FOR SEQ ID NO: 140:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 450 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 68...346
- (D) OTHER INFORMATION:

216

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 140:

```

TGGCGCTGTC ACCGAGGAAC CTGTCAATGT CGTCGAGCAG TACTGAACCG TTCCGAGAAA      60
GGCCAGC ATG AAC GTC ACC GTA TCC ATT CCG ACC ATC CTG CCG CCC CAC      109
Met Asn Val Thr Val Ser Ile Pro Thr Ile Leu Arg Pro His
      1              5              10
ACC GGC GGC CAG AAG AGT GTC TCG GCC AGC GGC GAT ACC TTG GGT GCC      157
Thr Gly Gly Gln Lys Ser Val Ser Ala Ser Gly Asp Thr Leu Gly Ala
      15              20              25              30
GTC ATC AGC GAC CTG GAG GCC AAC TAT TCG GGC ATT TCC GAG CGC CTG      205
Val Ile Ser Asp Leu Glu Ala Asn Tyr Ser Gly Ile Ser Glu Arg Leu
      35              40              45
ATG GAC CCG TCT TCC CCA GGT AAG TTG CAC CGC TTC GTG AAC ATC TAC      253
Met Asp Pro Ser Ser Pro Gly Lys Leu His Arg Phe Val Asn Ile Tyr
      50              55              60
GTC AAC GAC GAG GAC GTG CCG TTC TCC GGC GGC TTG GCC ACC GCG ATC      301
Val Asn Asp Glu Asp Val Arg Phe Ser Gly Gly Leu Ala Thr Ala Ile
      65              70              75
GGT GAC GGT GAC TCG GTC ACC ATC CTC CCC GCC GTG GCC GGT GCG TGAGC      351
Ala Asp Gly Asp Ser Val Thr Ile Leu Pro Ala Val Ala Gly Gly
      80              85              90
GGAGCACATG ACACGATACG ACTCGCTGTT GCAGGCTTGG GGCAACACGC CGCTGGTTGG      411
CCTGCAGCGA TTGTGCCCAC GCTGGGATGA CCGGCGAGA      450

```

(3) INFORMATION FOR SEQ ID NO: 141:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 93 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 141:

```

Met Asn Val Thr Val Ser Ile Pro Thr Ile Leu Arg Pro His Thr Gly
      1              5              10              15
Gly Gln Lys Ser Val Ser Ala Ser Gly Asp Thr Leu Gly Ala Val Ile
      20              25              30
Ser Asp Leu Glu Ala Asn Tyr Ser Gly Ile Ser Glu Arg Leu Met Asp
      35              40              45
Pro Ser Ser Pro Gly Lys Leu His Arg Phe Val Asn Ile Tyr Val Asn
      50              55              60

```

217

Asp Glu Asp Val Arg Phe Ser Gly Gly Leu Ala Thr Ala Ile Ala Asp
 65 70 75 80

Gly Asp Ser Val Thr Ile Leu Pro Ala Val Ala Gly Gly
 85 90

(3) INFORMATION FOR SEQ ID NO: 142:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 480 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 88...381
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 142:

GGTGTTCCTCCG CGGCCGGCTA TGACAACAGT CATTGTGCAT GACAAGTTAC AGGTATTAGG 60

TCCAGGTTCA ACAAGGAGAC AGGCAAC ATG GCA ACA CGT TTT ATG ACG GAT CCG 114
 Met Ala Thr Arg Phe Met Thr Asp Pro
 1 5

CAC GCG ATG CCG GAC ATG GCG GGC CGT TTT GAG GTG CAC GCC CAG ACG 162
 His Ala Met Arg Asp Met Ala Gly Arg Phe Glu Val His Ala Gln Thr
 10 15 20 25

CTG GAG GAC GAG GCT CGC CCG ATG TGG GCG TCC CCG CAA AAC ATC TCG 210
 Val Glu Asp Glu Ala Arg Arg Met Trp Ala Ser Ala Gln Asn Ile Ser
 30 35 40

GGC GCG GGC TGG AGT GGC ATG GCC GAG GCG ACC TCG CTA GAC ACC ATG 258
 Gly Ala Gly Trp Ser Gly Met Ala Glu Ala Thr Ser Leu Asp Thr Met
 45 50 55

GCC CAG ATG AAT CAG GCG TTT CGC AAC ATC GTG AAC ATG CTG CAC GGG 306
 Ala Gln Met Asn Gln Ala Phe Arg Asn Ile Val Asn Met Leu His Gly
 60 65 70

GTG CGT GAC GGG CTG GTT CGC GAC GCC AAC AAC TAC GAG CAG CAA GAG 354
 Val Arg Asp Gly Leu Val Arg Asp Ala Asn Asn Tyr Glu Gln Gln Glu
 75 80 85

CAG GCC TCC CAG CAG ATC CTC AGC AGC TAACCTCAGC CGCTCAGCA CAATACT 408
 Gln Ala Ser Gln Gln Ile Leu Ser Ser
 90 95

TTTACAAGCG AAGGAGAACA GGTTCGATGA CCATCAACTA TCAGTTCGGT GATGTCCAGC 468

CTCATGGCGC CA 480

218

(2) INFORMATION FOR SEQ ID NO: 143:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 99 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 143:

```

Met Ala Thr Arg Phe Met Thr Asp Pro His Ala Met Arg Asp Met Ala
 1             5             10             15

Gly Arg Phe Glu Val His Ala Gln Thr Val Glu Asp Glu Ala Arg Arg
          20             25             30

Met Trp Ala Ser Ala Gln Asn Ile Ser Gly Ala Gly Trp Ser Gly Met
          35             40             45

Ala Glu Ala Thr Ser Leu Asp Thr Met Ala Gln Met Asn Gln Ala Phe
          50             55             60

Arg Asn Ile Val Asn Met Leu His Gly Val Arg Asp Gly Leu Val Arg
          65             70             75             80

Asp Ala Asn Asn Tyr Glu Gln Gln Glu Gln Ala Ser Gln Gln Ile Leu
          85             90             95

Ser Ser

```

(3) INFORMATION FOR SEQ ID NO: 144:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 948 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 86...868
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 144:

```

GCCCCAGTCC TCGATCGCCT CATCGCCTTC ACCGGCGCGCC AGCCGACCGC AGGCCACGTG      60

TCCGCCACCT AACGAAGGA TGATC ATG CCC AAG AGA AGC GAA TAC AGG CAA      112
                Met Pro Lys Arg Ser Glu Tyr Arg Gln
                1             5

GGC ACG CCG AAC TGG GTC GAC CTT CAG ACC ACC GAT CAG TCC GCC GCC      160

```

219

Gly	Thr	Pro	Asn	Trp	Val	Asp	Leu	Gln	Thr	Thr	Asp	Gln	Ser	Ala	Ala		
10					15				20						25		
AAA	AAG	TTT	TAC	ACA	TCG	TTG	TTC	GGC	TGG	GGT	TAC	GAC	GAC	AAC	CCG		208
Lys	Lys	Phe	Tyr	Thr	Ser	Leu	Phe	Gly	Trp	Gly	Tyr	Asp	Asp	Asn	Pro		
				30					35					40			
GTC	CCC	GGA	GGC	GGT	GGG	GTC	TAT	TCC	ATG	GCC	ACG	CTG	AAC	GGC	GAA		256
Val	Pro	Gly	Gly	Gly	Gly	Val	Tyr	Ser	Met	Ala	Thr	Leu	Asn	Gly	Glu		
			45					50					55				
GCC	GTG	GCC	GCC	ATC	GCA	CCG	ATG	CCC	CCG	GGT	GCA	CCG	GAG	GGG	ATG		304
Ala	Val	Ala	Ala	Ile	Ala	Pro	Met	Pro	Pro	Gly	Ala	Pro	Glu	Gly	Met		
		60					65					70					
CCG	CCG	ATC	TGG	AAC	ACC	TAT	ATC	GCG	GTG	GAC	GAC	GTG	GAT	GCG	GTG		352
Pro	Pro	Ile	Trp	Asn	Thr	Tyr	Ile	Ala	Val	Asp	Asp	Val	Asp	Ala	Val		
		75				80					85						
GTG	GAC	AAG	GTG	GTG	CCC	GGG	GGC	GGG	CAG	GTG	ATG	ATG	CCG	GCC	TTC		400
Val	Asp	Lys	Val	Val	Pro	Gly	Gly	Gly	Gln	Val	Met	Met	Pro	Ala	Phe		
	90				95				100						105		
GAC	ATC	GGC	GAT	GCC	GGC	CCG	ATG	TGG	TTC	ATC	ACC	GAT	CCG	ACC	GGC		448
Asp	Ile	Gly	Asp	Ala	Gly	Arg	Met	Ser	Phe	Ile	Thr	Asp	Pro	Thr	Gly		
			110						115					120			
GCT	GCC	GTG	GGC	CTA	TGG	CAG	GCC	AAT	CGG	CAC	ATC	GGA	CGG	ACG	TTG		496
Ala	Ala	Val	Gly	Leu	Trp	Gln	Ala	Asn	Arg	His	Ile	Gly	Ala	Thr	Leu		
			125					130				135					
GTG	AAC	GAG	ACG	GGC	ACG	CTC	ATC	TGG	AAC	GAA	CTG	CTC	ACG	GAC	AAG		544
Val	Asn	Glu	Thr	Gly	Thr	Leu	Ile	Trp	Asn	Glu	Leu	Leu	Thr	Asp	Lys		
		140				145					150						
CCG	GAT	TTG	GCG	CTA	GCG	TTC	TAC	GAG	GCT	GTG	GTT	GSC	CTC	ACC	CAC		592
Pro	Asp	Leu	Ala	Leu	Ala	Phe	Tyr	Glu	Ala	Val	Val	Gly	Leu	Thr	His		
		155				160					165						
TGG	AGC	ATG	GAG	ATA	GCT	CCG	GGC	CAG	AAC	TAT	CCG	GTG	CTC	AAG	GCC		640
Ser	Ser	Met	Glu	Ile	Ala	Ala	Gly	Gln	Asn	Tyr	Arg	Val	Leu	Lys	Ala		
	170				175					180					185		
GGC	GAC	GCG	GAA	GTC	GGC	GGC	TGT	ATG	GAA	CCG	CCG	ATG	CCC	GGC	GTG		688
Gly	Asp	Ala	Glu	Val	Gly	Gly	Cys	Met	Glu	Pro	Pro	Met	Pro	Gly	Val		
			190					195				200					
CCG	AAT	CAT	TGG	CAC	GTC	TAC	TTT	GCG	GTG	GAT	GAC	GCC	GAC	GCC	ACG		736
Pro	Asn	His	Trp	His	Val	Tyr	Phe	Ala	Val	Asp	Asp	Ala	Asp	Ala	Thr		
		205						210				215					
GCG	GCC	AAA	GCC	GCC	GCA	GCG	GGC	GGC	CAG	GTG	ATT	GCG	GAA	CCG	GCT		784
Ala	Ala	Lys	Ala	Ala	Ala	Ala	Gly	Gly	Gln	Val	Ile	Ala	Glu	Pro	Ala		
		220					225				230						
GAC	ATT	CCG	TGG	GTG	GGC	CCG	TTC	GCC	GTG	TTG	TCC	GAT	CCG	CAG	GGC		832
Asp	Ile	Pro	Ser	Val	Gly	Arg	Phe	Ala	Val	Leu	Ser	Asp	Pro	Gln	Gly		
	235					240					245						

220

GCG ATC TTC AGT GTG TTG AAG CCC GCA CCG CAG CAA TAGGGAGCAT CCGGG 884
 Ala Ile Phe Ser Val Leu Lys Pro Ala Pro Gln Gln
 250 255 260

CAGGCCCGCC GCGCGGCAGA TTGGGAGAAT GCTAGAAGCT GCGCGCGGCG CCGCCC 940

(2) INFORMATION FOR SEQ ID NO: 145:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 261 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 145:

Met	Pro	Lys	Arg	Ser	Glu	Tyr	Arg	Gln	Gly	Thr	Pro	Asn	Trp	Val	Asp	1	5	10	15
Leu	Gln	Thr	Thr	Asp	Gln	Ser	Ala	Ala	Lys	Lys	Phe	Tyr	Thr	Ser	Leu	20	25	30	
Phe	Gly	Trp	Gly	Tyr	Asp	Asp	Asn	Pro	Val	Pro	Gly	Gly	Gly	Gly	Val	35	40	45	
Tyr	Ser	Met	Ala	Thr	Leu	Asn	Gly	Glu	Ala	Val	Ala	Ala	Ile	Ala	Pro	50	55	60	
Met	Pro	Pro	Gly	Ala	Pro	Glu	Gly	Met	Pro	Pro	Ile	Trp	Asn	Thr	Tyr	65	70	75	80
Ile	Ala	Val	Asp	Asp	Val	Asp	Ala	Val	Val	Asp	Lys	Val	Val	Pro	Gly	85	90	95	
Gly	Gly	Gln	Val	Met	Met	Pro	Ala	Phe	Asp	Ile	Gly	Asp	Ala	Gly	Arg	100	105	110	
Met	Ser	Phe	Ile	Thr	Asp	Pro	Thr	Gly	Ala	Ala	Val	Gly	Leu	Trp	Gln	115	120	125	
Ala	Asn	Arg	His	Ile	Gly	Ala	Thr	Leu	Val	Asn	Glu	Thr	Gly	Thr	Leu	130	135	140	
Ile	Trp	Asn	Glu	Leu	Leu	Thr	Asp	Lys	Pro	Asp	Leu	Ala	Leu	Ala	Phe	145	150	155	160
Tyr	Glu	Ala	Val	Val	Gly	Leu	Thr	His	Ser	Ser	Met	Glu	Ile	Ala	Ala	165	170	175	
Gly	Gln	Asn	Tyr	Arg	Val	Leu	Lys	Ala	Gly	Asp	Ala	Glu	Val	Gly	Gly	180	185	190	
Cys	Met	Glu	Pro	Pro	Met	Pro	Gly	Val	Pro	Asn	His	Trp	His	Val	Tyr	195	200	205	

221

Phe Ala Val Asp Asp Ala Asp Ala Thr Ala Ala Lys Ala Ala Ala Ala
 210 215 220

Gly Gly Gln Val Ile Ala Glu Pro Ala Asp Ile Pro Ser Val Gly Arg
 225 230 235 240

Phe Ala Val Leu Ser Asp Pro Gln Gly Ala Ile Phe Ser Val Leu Lys
 245 250 255

Pro Ala Pro Gln Gln
 260

(2) INFORMATION FOR SEQ ID NO: 146:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 280 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 47...247
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 146:

CCGAAGGGCG GTGCACCGCA CCCAGAAGAA AAGGAAGAT CGAGAA ATG CCA CAG 55
 Met Pro Gln
 1

GGA ACT GTG AAG TGG TTC AAC GCG GAG AAG GGG TTC GGC TTT ATC GCC 103
 Gly Thr Val Lys Trp Phe Asn Ala Glu Lys Gly Phe Gly Phe Ile Ala
 5 10 15

CCC GAA GAC GGT TCC GCG GAT GTA TTT GTC CAC TAC ACG GAG ATC CAG 151
 Pro Glu Asp Gly Ser Ala Asp Val Phe Val His Tyr Thr Glu Ile Gln
 20 25 30 35

GGA ACG GGC TTC CGC ACC CTT GAA GAA AAC CAG AAG GTC GAG TTC GAG 199
 Gly Thr Gly Phe Arg Thr Leu Glu Glu Asn Gln Lys Val Glu Phe Glu
 40 45 50

ATC GGC CAC AGC CCT AAG GGC CCC CAG GCC ACC GGA GTC CGC TCG CTC T 248
 Ile Gly His Ser Pro Lys Gly Pro Gln Ala Thr Gly Val Arg Ser Leu
 55 60 65

GAGTACCCC CGCGAGCAGA CGCAAAAAGC CC 280

(2) INFORMATION FOR SEQ ID NO: 147:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 67 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

222

- (ii) MOLECULE TYPE: protein
 (v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 147:

```
Met Pro Gln Gly Thr Val Lys Trp Phe Asn Ala Glu Lys Gly Phe Gly
 1             5             10             15
Phe Ile Ala Pro Glu Asp Gly Ser Ala Asp Val Phe Val His Tyr Thr
          20             25             30
Glu Ile Gln Gly Thr Gly Phe Arg Thr Leu Glu Glu Asn Gln Lys Val
          35             40             45
Glu Phe Glu Ile Gly His Ser Pro Lys Gly Pro Gln Ala Thr Gly Val
          50             55             60
Arg Ser Leu
65
```

(2) INFORMATION FOR SEQ ID NO: 148:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 540 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
 (B) LOCATION: 105...491
 (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 148:

```
ATCGTGTCTCT ATCGAGAACC CCGGCCGGTA TCASAAACGCG CCAGAGCGCA AACCTTTATA      60
ACTTCGTGTTC CCAAAATGTGA CGACCATGGA CCAAGGTTCC TGAG ATG AAC CTA CCG      116
                                     Met Asn Leu Arg
                                     1
CGC CAT CAG ACC CTG ACG CTG CGA CTG CTG GCG GCA TCC GCG GGC ATT      164
Arg His Gln Thr Leu Thr Leu Arg Leu Leu Ala Ala Ser Ala Gly Ile
 5             10             15             20
CTC AGC GCC GCG GCC TTC GCC GCG CCA GCA CAG GCA AAC CCC GTC GAC      212
Leu Ser Ala Ala Ala Phe Ala Ala Pro Ala Gln Ala Asn Pro Val Asp
          25             30             35
GAC GCG TTC ATC GCC GCG CTG AAC AAT GCC GCG GTC AAC TAC GGC GAT      260
Asp Ala Phe Ile Ala Ala Leu Asn Asn Ala Gly Val Asn Tyr Gly Asp
          40             45             50
CCG GTC GAC GCC AAA GCG CTG GGT CAG TCC GTC TGC CCG ATC CTG GCC      308
Pro Val Asp Ala Lys Ala Leu Gly Gln Ser Val Cys Pro Ile Leu Ala
          55             60             65
```


224

(2) INFORMATION FOR SEQ ID NO: 150:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 25...354
- (D) OTHER INFORMATION:

(ix) FEATURE:

- (A) NAME/KEY: mat_peptide
- (B) LOCATION: 109...357

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 150:

ATAGTTTGGG GAAGGTGTCC ATAA ATG AGG CTS TCG TTG ACC GCA TTG AGC	51
Met Arg Leu Ser Leu Thr Ala Leu Ser	
-28 -25 -20	
GCC GGT GTA GCC GCC GTG GCA ATG TCG TTG ACC GTC GGG GCC GGG GTC	99
Ala Gly Val Gly Ala Val Ala Met Ser Leu Thr Val Gly Ala Gly Val	
-15 -10 -5	
GCC TCC GCA GAT CCC GTG GAC GCG GTC ATT AAC ACC ACC TGC AAT TAC	147
Ala Ser Ala Asp Pro Val Asp Ala Val Ile Asn Thr Thr Cys Asn Tyr	
1 5 10	
GGG CAG GTA GTA GCT GCG CTC AAC GCG ACG GAT CCG GGG GCT GCC GCA	195
Gly Gln Val Val Ala Ala Leu Asn Ala Thr Asp Pro Gly Ala Ala Ala	
15 20 25	
CAG TTC AAC GCC TCA CCG GTG GCG CAG TCC TAT TTG CGC AAT TTC CTC	243
Gln Phe Asn Ala Ser Pro Val Ala Gln Ser Tyr Leu Arg Asn Phe Leu	
30 35 40 45	
GCC GCA CCG CCA CCT CAG CGC GCT GCC ATG GCC GCG CAA TTG CAA GCT	291
Ala Ala Pro Pro Pro Gln Arg Ala Ala Met Ala Ala Gln Leu Gln Ala	
50 55 60	
GTG CCG GGG GCG GCA CAG TAC ATC GGC CTT GTC GAG TCG GTT GCC GGC	339
Val Pro Gly Ala Ala Gln Tyr Ile Gly Leu Val Glu Ser Val Ala Gly	
65 70 75	
TCC TGC AAC AAC TAT TAAGCCCATG CCGGCCCAT CCGCGACCC GGCATCCTCG	394
Ser Cys Asn Asn Tyr	
80	
CCGGGG	400

(2) INFORMATION FOR SEQ ID NO: 151:

225

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 110 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 151:

```

Met Arg Leu Ser Leu Thr Ala Leu Ser Ala Gly Val Gly Ala Val Ala
-28          -25          -20          -15

Met Ser Leu Thr Val Gly Ala Gly Val Ala Ser Ala Asp Pro Val Asp
          -10          -5          1

Ala Val Ile Asn Thr Thr Cys Asn Tyr Gly Gln Val Val Ala Ala Leu
5          10          15          20

Asn Ala Thr Asp Pro Gly Ala Ala Ala Gln Phe Asn Ala Ser Pro Val
          25          30          35

Ala Gln Ser Tyr Leu Arg Asn Phe Leu Ala Ala Pro Pro Pro Gln Arg
          40          45          50

Ala Ala Met Ala Ala Gln Leu Gln Ala Val Pro Gly Ala Ala Gln Tyr
          55          60          65

Ile Gly Leu Val Glu Ser Val Ala Gly Ser Cys Asn Asn Tyr
          70          75          80

```

(2) INFORMATION FOR SEQ ID NO: 152:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 990 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 93...890
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 152:

```

AATAGTAAIA TCCTGTGCG GTTGCAAAAC GTGTACCGA GGTCCGCGAG TCGAGCGCTG      60
CGGGCCGCCT TCGAGGAGGA CCAACCACAG TC ATG ACG AAC ATC CTG CTC CTG      113
                               Met Thr Asn Ile Val Val Leu
                               1          5

ATC AAG CAG CTC CCA GAT ACC TGG TCG GAG CGC AAG CTG ACC GAC GGC      161
Ile Lys Gln Val Pro Asp Thr Trp Ser Glu Arg Lys Leu Thr Asp Gly
          10          15          20

GAT TTC ACG CTG GAC CGC GAG GCC GCC GAC GCG GTG CTG GAC GAG ATC      209

```

226

Asp	Phe	Thr	Leu	Asp	Arg	Glu	Ala	Ala	Asp	Ala	Val	Leu	Asp	Glu	Ile	
25						30					35					
AAC	GAG	CGC	GCC	GTG	GAG	GAA	GGC	CTA	CAG	ATT	CGG	GAG	AAA	GAG	GCC	257
Asn	Glu	Arg	Ala	Val	Glu	Glu	Ala	Leu	Gln	Ile	Arg	Glu	Lys	Glu	Ala	
40					45				50					55		
GCC	GAC	GGC	ATC	GAA	GGG	TCG	GTA	ACC	GTG	CTG	ACG	GGC	GGC	CCC	GAG	305
Ala	Asp	Gly	Ile	Glu	Gly	Ser	Val	Thr	Val	Leu	Thr	Ala	Gly	Pro	Glu	
			60					65						70		
CGC	GCC	ACC	GAG	GGC	ATC	CGC	AAG	GGC	CTG	TCG	ATG	GGT	GCC	GAC	AAG	353
Arg	Ala	Thr	Glu	Ala	Ile	Arg	Lys	Ala	Leu	Ser	Met	Gly	Ala	Asp	Lys	
			75				80						85			
GCC	GTC	CAC	CTA	AAG	GAC	GAC	GGC	ATG	CAC	GGC	TCG	GAC	GTC	ATC	CAA	401
Ala	Val	His	Leu	Lys	Asp	Asp	Gly	Met	His	Gly	Ser	Asp	Val	Ile	Gln	
		90					95					100				
ACC	GGG	TGG	GCT	TTG	GCG	CGC	GGC	TTG	GGC	ACC	ATC	GAG	GGC	ACC	GAG	449
Thr	Gly	Trp	Ala	Leu	Ala	Arg	Ala	Leu	Gly	Thr	Ile	Glu	Gly	Thr	Glu	
105					110						115					
CTG	GTG	ATC	GCA	GGC	AAC	GAA	TCG	ACC	GAC	GGG	GTG	GGC	GGT	GGG	GTG	497
Leu	Val	Ile	Ala	Gly	Asn	Glu	Ser	Thr	Asp	Gly	Val	Gly	Gly	Ala	Val	
120					125					130					135	
CCG	GCC	ATC	ATC	GCC	GAG	TAC	CTG	GGC	CTG	CCG	CAG	CTC	ACC	CAC	CTG	545
Pro	Ala	Ile	Ile	Ala	Glu	Tyr	Leu	Gly	Leu	Pro	Gln	Leu	Thr	His	Leu	
			140						145					150		
CGC	AAA	GTG	TCG	ATC	GAG	GGC	GGC	AAG	ATC	ACC	GGC	GAG	CGT	GAG	ACC	593
Arg	Lys	Val	Ser	Ile	Glu	Gly	Gly	Lys	Ile	Thr	Gly	Glu	Arg	Glu	Thr	
		155					160						165			
GAT	GAG	GGC	GTA	TTC	ACC	CTC	GAG	GGC	ACG	CTG	CCG	GGG	GTG	ATC	AGC	641
Asp	Glu	Gly	Val	Phe	Thr	Leu	Glu	Ala	Thr	Leu	Pro	Ala	Val	Ile	Ser	
		170					175						180			
GTG	AAC	GAG	AAG	ATC	AAC	GAG	CCG	CGC	TTC	CCG	TCC	TTC	AAA	GUC	ATC	689
Val	Asn	Glu	Lys	Ile	Asn	Glu	Pro	Arg	Phe	Pro	Ser	Phe	Lys	Gly	Ile	
		185				190					195					
ATG	GCC	GCC	AAG	AAG	AAG	GAA	GTT	ACC	GTG	CTG	ACC	CTG	GCC	GAG	ATC	737
Met	Ala	Ala	Lys	Lys	Lys	Glu	Val	Thr	Val	Leu	Thr	Leu	Ala	Glu	Ile	
200					205					210				215		
GGT	GTC	GAG	AGC	GAC	GAG	GTG	GGG	CTG	GGC	AAC	GCC	GGA	TCC	ACC	GTG	785
Gly	Val	Glu	Ser	Asp	Glu	Val	Gly	Leu	Ala	Asn	Ala	Gly	Ser	Thr	Val	
			220						225					230		
CTG	GGC	TCG	ACG	CCC	AAA	CCG	GGC	AAG	ACT	GCC	GGG	GAG	AAG	GTC	ACC	833
Leu	Ala	Ser	Thr	Pro	Lys	Pro	Ala	Lys	Thr	Ala	Gly	Glu	Lys	Val	Thr	
		235					240						245			
GAC	GAG	GGT	GAA	GCC	GGC	AAC	CAG	ATC	GTG	CAG	TAC	CTG	GTT	GCC	CAG	881
Asp	Glu	Gly	Glu	Gly	Gly	Asn	Gln	Ile	Val	Gln	Tyr	Leu	Val	Ala	Gln	
		250					255						260			

227

AAA ATC ATC TAAGACATAC GCACCTCCCA AAGACGAGAG CGATATAACC CATGGCTGA 939
 Lys Ile Ile
 265

AGTACTGGTG CTCGTTGAGC ACGCTGAAGG CGCSTTAAAG AAGGTCAGCG C 990

(2) INFORMATION FOR SEQ ID NO: 153:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 266 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 153:

Met	Thr	Asn	Ile	Val	Val	Leu	Ile	Lys	Gln	Val	Pro	Asp	Thr	Trp	Ser	1	5	10	15
Glu	Arg	Lys	Leu	Thr	Asp	Gly	Asp	Phe	Thr	Leu	Asp	Arg	Glu	Ala	Ala	20	25	30	
Asp	Ala	Val	Leu	Asp	Glu	Ile	Asn	Glu	Arg	Ala	Val	Glu	Glu	Ala	Leu	35	40	45	
Gln	Ile	Arg	Glu	Lys	Glu	Ala	Ala	Asp	Gly	Ile	Glu	Gly	Ser	Val	Thr	50	55	60	
Val	Leu	Thr	Ala	Gly	Pro	Glu	Arg	Ala	Thr	Glu	Ala	Ile	Arg	Lys	Ala	65	70	75	80
Leu	Ser	Met	Gly	Ala	Asp	Lys	Ala	Val	His	Leu	Lys	Asp	Asp	Gly	Met	85	90	95	
His	Gly	Ser	Asp	Val	Ile	Gln	Thr	Gly	Trp	Ala	Leu	Ala	Arg	Ala	Leu	100	105	110	
Gly	Thr	Ile	Glu	Gly	Thr	Glu	Leu	Val	Ile	Ala	Gly	Asn	Glu	Ser	Thr	115	120	125	
Asp	Gly	Val	Gly	Gly	Ala	Val	Pro	Ala	Ile	Ile	Ala	Glu	Tyr	Leu	Gly	130	135	140	
Leu	Pro	Gln	Leu	Thr	His	Leu	Arg	Lys	Val	Ser	Ile	Glu	Gly	Gly	Lys	145	150	155	160
Ile	Thr	Gly	Glu	Arg	Glu	Thr	Asp	Glu	Gly	Val	Phe	Thr	Leu	Glu	Ala	165	170	175	
Thr	Leu	Pro	Ala	Val	Ile	Ser	Val	Asn	Glu	Lys	Ile	Asn	Glu	Pro	Arg	180	185	190	
Phe	Pro	Ser	Phe	Lys	Gly	Ile	Met	Ala	Ala	Lys	Lys	Lys	Glu	Val	Thr	195	200	205	
Val	Leu	Thr	Leu	Ala	Glu	Ile	Gly	Val	Glu	Ser	Asp	Glu	Val	Gly	Leu				

228

210	215	220
Ala Asn Ala Gly Ser Thr Val Leu Ala Ser Thr Pro Lys Pro Ala Lys		
225	230	235
Thr Ala Gly Glu Lys Val Thr Asp Glu Gly Glu Gly Gly Asn Gln Ile		
245	250	255
Val Gln Tyr Leu Val Ala Gln Lys Ile Ile		
260	265	

(2) INFORMATION FOR SEQ ID NO: 154:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 25 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 154:

CTGAGATCTA TGAACCTACG GCGCC 25

(2) INFORMATION FOR SEQ ID NO: 155:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 35 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 155:

CTCCCATGGT ACCCTAGGAC CCGGGCAGCC CCGGC 35

(2) INFORMATION FOR SEQ ID NO: 156:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 29 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 156:

CTGAGATCTA TGAGGCTGTC GTTGACCGC 29

(2) INFORMATION FOR SEQ ID NO: 157:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 157:

CTCCCCGGGC TTANTAGTTG TTGCAGGAGC

30

(2) INFORMATION FOR SEQ ID NO: 158:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 33 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 158:

GCTTAGATCT ATGATTTTCT GGGCAACCAG GTA

33

(2) INFORMATION FOR SEQ ID NO: 159:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 159:

GCTTCCATGG GCGAGGCACA GCGGTGGGAA

30

(2) INFORMATION FOR SEQ ID NO: 160:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 160:

CTGAGATCTA GAATGCCACA GGGAACTGTG

30

(2) INFORMATION FOR SEQ ID NO: 161:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 161:

TCTCCCCGGG GTAACTCAGA GCGAGCGGAC

30

(2) INFORMATION FOR SEQ ID NO: 162:

230

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 162:

CTGAGATCTA TGAACGTCAC CHTATCC

37

(2) INFORMATION FOR SEQ ID NO: 163:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 163:

TCTCCCGGGG CTCACCCACC GGGCACC

27

(2) INFORMATION FOR SEQ ID NO: 164:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 164:

CTGAGATCTA TGGCAACACC TTTTATGACG

30

(2) INFORMATION FOR SEQ ID NO: 165:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 165:

CTCCCCGGGT TAGCTGCTGA GGATCTGCTN

30

(2) INFORMATION FOR SEQ ID NO: 166:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

231

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 166:

CTGAAGATCT ATGCCCAAGA GAAGCGAATA C

31

(2) INFORMATION FOR SEQ ID NO: 167:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 167:

CGGCAGCTGC TAGCATTCTC CGAATCTGCC G

31

(2) INFORMATION FOR SEQ ID NO: 168:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 168:

Pro	Gln	Gly	Thr	Val	Lys	Trp	Phe	Asn	Ala	Glu	Lys	Gly	Phe	Gly
1				5					10					15

(2) INFORMATION FOR SEQ ID NO: 169:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(ix) FEATURE:

- (A) NAME/KEY: Other
- (B) LOCATION: 15
- (D) OTHER INFORMATION: Xaa is unknown

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 169:

Asn	Val	Thr	Val	Ser	Ile	Pro	Thr	Ile	Leu	Arg	Pro	Xaa	Xaa	Xaa
1				5					10					15

(2) INFORMATION FOR SEQ ID NO: 170:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single

232

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(ix) FEATURE:

(A) NAME/KEY: Other

(B) LOCATION: 1

(D) OTHER INFORMATION: Thr Could also be Ala

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 170:

Thr	Arg	Phe	Met	Thr	Asp	Pro	His	Ala	Met	Arg	Asp	Met	Ala	Gly
1				5					10					15

(2) INFORMATION FOR SEQ ID NO: 171:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 15 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 171:

Pro	Lys	Arg	Ser	Glu	Tyr	Arg	Gln	Gly	Thr	Pro	Asn	Trp	Val	Asp
1		5		10				15						

(2) INFORMATION FOR SEQ ID NO:172:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 404 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:172:

Met	Ala	Thr	Val	Asn	Arg	Ser	Arg	His	His	His	His	His	His	His	His
1				5				10						15	
Ile	Glu	Gly	Arg	Ser	Phe	Ser	Arg	Pro	Gly	Leu	Pro	Val	Glu	Tyr	Leu
			20					25					30		
Gln	Val	Pro	Ser	Pro	Ser	Met	Gly	Arg	Asp	Ile	Lys	Val	Gln	Phe	Gln
		35				40					45				
Ser	Gly	Gly	Asn	Asn	Ser	Pro	Ala	Val	Tyr	Leu	Leu	Asp	Gly	Leu	Arg
		50				55				60					
Ala	Gln	Asp	Asp	Tyr	Asn	Gly	Trp	Asp	Ile	Asn	Thr	Pro	Ala	Phe	Glu
65				70				75						80	
Trp	Tyr	Tyr	Gln	Ser	Gly	Leu	Ser	Ile	Val	Met	Pro	Val	Gly	Gly	Gln
			85					90					95		
Ser	Ser	Phe	Tyr	Ser	Asp	Trp	Tyr	Ser	Pro	Ala	Cys	Gly	Lys	Ala	Gly
		100				105						110			
Cys	Gln	Thr	Tyr	Lys	Trp	Glu	Thr	Phe	Leu	Thr	Ser	Glu	Leu	Pro	Gln
		115				120					125				
Trp	Leu	Ser	Ala	Asn	Arg	Ala	Val	Lys	Pro	Thr	Gly	Ser	Ala	Ala	Ile

233

130		135		140
Gly Leu Ser Met Ala	Gly Ser Ser Ala Met	Ile Leu Ala Ala Tyr His		
145	150	155		160
Pro Gln Gln Phe Ile	Tyr Ala Gly Ser Leu Ser Ala Leu Leu Asp Pro			
	165	170		175
Ser Gln Gly Met Gly	Pro Ser Leu Ile Gly Leu Ala Met Gly Asp Ala			
	180	185		190
Gly Gly Tyr Lys Ala Ala	Asp Met Trp Gly Pro Ser Ser Asp Pro Ala			
	195	200		205
Trp Glu Arg Asn Asp	Pro Thr Gln Gln Ile Pro Lys Leu Val Ala Asn			
	210	215		220
Asn Thr Arg Leu Trp	Val Tyr Cys Gly Asn Gly Thr Pro Asn Glu Leu			
225	230	235		240
Gly Gly Ala Asn Ile	Pro Ala Glu Phe Leu Glu Asn Phe Val Arg Ser			
	245	250		255
Ser Asn Leu Lys Phe	Gln Asp Ala Tyr Asn Ala Ala Gly Gly His Asn			
	260	265		270
Ala Val Phe Asn Phe	Pro Pro Asn Gly Thr His Ser Trp Glu Tyr Trp			
	275	280		285
Gly Ala Gln Leu Asn	Ala Met Lys Gly Asp Leu Gln Ser Ser Leu Gly			
	290	295		300
Ala Gly Lys Leu Ala	Met Thr Glu Gln Gln Trp Asn Phe Ala Gly Ile			
305	310	315		320
Glu Ala Ala Ala Ser	Ala Ile Gln Gly Asn Val Thr Ser Ile His Ser			
	325	330		335
Leu Leu Asp Glu Gly	Lys Gln Ser Leu Thr Lys Leu Ala Ala Ala Trp			
	340	345		350
Gly Gly Ser Gly Ser	Glu Ala Tyr Gln Gly Val Gln Gln Lys Trp Asp			
	355	360		365
Ala Thr Ala Thr Glu	Leu Asn Asn Ala Leu Gln Asn Leu Ala Arg Thr			
	370	375		380
Ile Ser Glu Ala Gly	Gln Ala Met Ala Ser Thr Glu Gly Asn Val Thr			
385	390	395		400
Gly Met Phe Ala				

(2) INFORMATION FOR SEQ ID NO:173:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 403 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:173:

Met Ala Thr Val Asn Arg Ser Arg His His His His His His His	
1 5 10 15	
Ile Glu Gly Arg Ser Met Thr Glu Gln Gln Trp Asn Phe Ala Gly Ile	
20 25 30	
Glu Ala Ala Ala Ser Ala Ile Gln Gly Asn Val Thr Ser Ile His Ser	
35 40 45	
Leu Leu Asp Glu Gly Lys Gln Ser Leu Thr Lys Leu Ala Ala Ala Trp	
50 55 60	
Gly Gly Ser Gly Ser Glu Ala Tyr Gln Gly Val Gln Gln Lys Trp Asp	
65 70 75 80	
Ala Thr Ala Thr Glu Leu Asn Asn Ala Leu Gln Asn Leu Ala Arg Thr	

234

85					90					95					
Ile	Ser	Glu	Ala	Gly	Gln	Ala	Met	Ala	Ser	Thr	Glu	Gly	Asn	Val	Thr
100					105					110					
Gly	Met	Phe	Ala	Lys	Leu	Phe	Ser	Arg	Pro	Gly	Leu	Pro	Val	Glu	Tyr
115					120					125					
Leu	Gln	Val	Pro	Ser	Pro	Ser	Met	Gly	Arg	Asp	Ile	Lys	Val	Gln	Phe
130					135					140					
Gln	Ser	Gly	Gly	Asn	Asn	Ser	Pro	Ala	Val	Tyr	Leu	Leu	Asp	Gly	Leu
145					150					155					
Arg	Ala	Gln	Asp	Asp	Tyr	Asn	Gly	Trp	Asp	Ile	Asn	Thr	Pro	Ala	Phe
160					165					170					
Glu	Trp	Tyr	Tyr	Gln	Ser	Gly	Leu	Ser	Ile	Val	Met	Pro	Val	Gly	Gly
180					185					190					
Gln	Ser	Ser	Phe	Tyr	Ser	Asp	Trp	Tyr	Ser	Pro	Ala	Cys	Gly	Lys	Ala
195					200					205					
Gly	Cys	Gln	Thr	Tyr	Lys	Trp	Glu	Thr	Phe	Leu	Thr	Ser	Glu	Leu	Pro
210					215					220					
Gln	Trp	Leu	Ser	Ala	Asn	Arg	Ala	Val	Lys	Pro	Thr	Gly	Ser	Ala	Ala
225					230					235					
Ile	Gly	Leu	Ser	Met	Ala	Gly	Ser	Ser	Ala	Met	Ile	Leu	Ala	Ala	Tyr
240					245					250					
His	Pro	Gln	Gln	Phe	Ile	Tyr	Ala	Gly	Ser	Leu	Ser	Ala	Leu	Leu	Asp
255					260					265					
Pro	Ser	Gln	Gly	Met	Gly	Pro	Ser	Leu	Ile	Gly	Leu	Ala	Met	Gly	Asp
270					275					280					
Ala	Gly	Gly	Tyr	Lys	Ala	Ala	Asp	Met	Trp	Gly	Pro	Ser	Ser	Asp	Pro
285					290					295					
Ala	Trp	Glu	Arg	Asn	Asp	Pro	Thr	Gln	Gln	Ile	Pro	Lys	Leu	Val	Ala
300					305					310					
Asn	Asn	Thr	Arg	Leu	Trp	Val	Tyr	Cys	Gly	Asn	Gly	Thr	Pro	Asn	Glu
315					320					325					
Leu	Gly	Gly	Ala	Asn	Ile	Pro	Ala	Glu	Phe	Leu	Glu	Asn	Phe	Val	Arg
330					335					340					
Ser	Ser	Asn	Leu	Lys	Phe	Gln	Asp	Ala	Tyr	Asn	Ala	Ala	Gly	Gly	His
345					350					355					
Asn	Ala	Val	Phe	Asn	Phe	Pro	Pro	Asn	Gly	Thr	His	Ser	Trp	Glu	Tyr
360					365					370					
Trp	Gly	Ala	Gln	Leu	Asn	Ala	Met	Lys	Gly	Asp	Leu	Gln	Ser	Ser	Leu
375					380					385					
Gly	Ala	Gly													

CLAIMS

1. A substantially pure polypeptide fragment which

- a) comprises an amino acid sequence selected from the sequences shown in SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 17-23, 42, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72-86, 88, 90, 92, 94, 141, 143, 145, 147, 149, 151, 153, and 168-171,
- b) comprises a subsequence of the polypeptide fragment defined in a) which has a length of at least 6 amino acid residues, said subsequence being immunologically equivalent to the polypeptide defined in a) with respect to the ability of evoking a protective immune response against infections with mycobacteria belonging to the tuberculosis complex or with respect to the ability of eliciting a diagnostically significant immune response indicating previous or ongoing sensitization with antigens derived from mycobacteria belonging to the tuberculosis complex, or
- c) comprises an amino acid sequence having a sequence identity with the polypeptide defined in a) or the subsequence defined in b) of at least 70% and at the same time being immunologically equivalent to the polypeptide defined in a) with respect to the ability of evoking a protective immune response against infections with mycobacteria belonging to the tuberculosis complex or with respect to the ability of eliciting a diagnostically significant immune response indicating previous or ongoing sensitization with antigens derived from mycobacteria belonging to the tuberculosis complex,

with the proviso that

- i) the polypeptide fragment is in essentially pure form when consisting of the amino acid sequence 1-96 of SEQ ID NO: 2 or when consisting of the amino acid sequence 87-108 of SEQ ID NO: 4 fused to β -galactosidase,
- 5 ii) the degree of sequence identity in c) is at least 95% when the polypeptide comprises a homologue of a polypeptide which has the amino acid sequence SEQ ID NO: 12 or a subsequence thereof as defined in b), and
- 10 iii) the polypeptide fragment contains a threonine residue corresponding to position 213 in SEQ ID NO: 42 when comprising an amino acid sequence of at least 6 amino acids in SEQ ID NO: 42.
2. The polypeptide fragment according to claim 1 in essentially pure form.
- 15 3. The polypeptide fragment according to claim 1 or 2, which comprises an epitope for a T-helper cell.
4. The polypeptide fragment according to any of the preceding claims, which has a length of at least 7 amino acid residues, such as at least 8, at least 9, at least 10, at least 12, at
20 least 14, at least 16, at least 18, at least 20, at least 22, at least 24, and at least 30 amino acid residues.
5. The polypeptide fragment according to any of the preceding claims, which is free from amino acid residues -30 to -1 in SEQ ID NO: 6 and/or -32 to -1 in SEQ ID NO: 10 and/or -8 to
25 -1 in SEQ ID NO: 12 and/or -32 to -1 in SEQ ID NO: 14 and/or -33 to -1 in SEQ ID NO: 42 and/or -38 to -1 in SEQ ID NO: 52 and/or -33 to -1 in SEQ ID NO: 56 and/or -56 to -1 in SEQ ID NO: 58 and/or -28 to -1 in SEQ ID NO: 151.
6. The polypeptide fragment according to any of the preceding
30 claims which is free from any signal sequence.

7. The polypeptide fragment according to any of the preceding claims which

- 1) induces a release of IFN- γ from primed memory T-lymphocytes withdrawn from a mouse within 2 weeks of primary infection or within 4 days after the mouse has been re-challenge infected with mycobacteria belonging to the tuberculosis complex, the induction performed by the addition of the polypeptide to a suspension comprising about 200,000 spleen cells per ml, the addition of the polypeptide resulting in a concentration of 1-4 μ g polypeptide per ml suspension, the release of IFN- γ being assessable by determination of IFN- γ in supernatant harvested 2 days after the addition of the polypeptide to the suspension, and/or
- 2) induces a release of IFN- γ of at least 300 pg above background level from about 1000,000 human PBMC (peripheral blood mononuclear cells) per ml isolated from TB patients in the first phase of infection, or from healthy BCG vaccinated donors, or from healthy contacts to TB patients, the induction being performed by the addition of the polypeptide to a suspension comprising the about 1,000,000 PBMC per ml, the addition of the polypeptide resulting in a concentration of 1-4 μ g polypeptide per ml suspension, the release of IFN- γ being assessable by determination of IFN- γ in supernatant harvested 2 days after the addition of the polypeptide to the suspension; and/or
- 3) induces an IFN- γ release from bovine PBMC derived from animals previously sensitized with mycobacteria belonging to the tuberculosis complex, said release being at least two times the release observed from bovine PBMC derived from animals not previously sensitized with mycobacteria belonging to the tuberculosis complex.

8. A polypeptide fragment according to any of the preceding claims, wherein the sequence identity in c) is at least 80%, such as at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%,
5 at least 97%, at least 98%, at least 99%, and at least 99.5%.

9. A fusion polypeptide comprising at least one polypeptide fragment according to any of the preceding claims and at least one fusion partner.

10. A fusion polypeptide according to claim 56, wherein the
10 fusion partner is selected from the group consisting of a polypeptide fragment as defined in any of claims 1-8, and an other polypeptide fragment derived from a bacterium belonging to the tuberculosis complex, such as ESAT-6 or at least one T-cell epitope thereof, MPB64 or at least one T-cell epitope
15 thereof, MPT64 or at least one T-cell epitope thereof, and MPB59 or at least one T-cell epitope thereof.

11. A fusion polypeptide fragment which comprises

1) a first amino acid sequence including at least one
stretch of amino acids constituting a T-cell epitope
20 derived from the *M. tuberculosis* protein ESAT-6, and a second amino acid sequence including at least one T-cell epitope derived from a *M. tuberculosis* protein different from ESAT-6 and/or including a stretch of
25 amino acids which protects the first amino acid sequence from in vivo degradation or post-translational processing; or

2) a first amino acid sequence including at least one
stretch of amino acids constituting a T-cell epitope
30 derived from the *M. tuberculosis* protein MPT59, and a second amino acid sequence including at least one T-cell epitope derived from a *M. tuberculosis* protein different from MPT59 and/or including a stretch of amino acids which protects the first amino acid

sequence from *in vivo* degradation or post-translational processing.

12. A fusion polypeptide fragment according to claim 11,
wherein the first amino acid sequence is situated C-termi-
5 nally to the second amino acid sequence.

13. A fusion polypeptide fragment according to claim 11,
wherein the first amino acid sequence is situated N-termi-
nally to the second amino acid sequence.

14. A fusion polypeptide fragment according to any of claims
10 11-13, wherein the at least one T-cell epitope included in
the second amino acid sequence is derived from a *M. tubercu-*
losis polypeptide selected from the group consisting of a
polypeptide fragment according to any of claims 1-55, DnaK,
GroEL, urease, glutamine synthetase, the proline rich com-
plex, L-alanine dehydrogenase, phosphate binding protein, Ag
15 85 complex, HBHA (heparin binding hemagglutinin), MPT51,
MPT64, superoxide dismutase, 19 kDa lipoprotein, α -crystal-
lin, GroES, MPT59 when the first T-cell epitope is derived
from ESAT-6, and ESAT-6 when the first T-cell epitope is
20 derived from MPT59.

15. A fusion polypeptide fragment according to any of claims
11-14, wherein the first and second T-cell epitopes each have
a sequence identity of at least 70% with the natively occur-
ring sequence in the proteins from which they are derived.

25 16. A fusion polypeptide according to any of claims 11-15,
wherein the first and/or second amino acid sequence have a
sequence identity of at least 70% with the protein from which
they are derived.

17. A fusion polypeptide fragment according to any of claims
30 11-16, wherein the first amino acid sequence is the amino
acid sequence of ESAT-6 or of MPT59 and/or the second amino
acid sequence is the amino acid sequence of a *M. tuberculosis*

- polypeptide selected from the group consisting of a polypeptide fragment according to any of claims 1-8, DnaK, GroEL, urease, glutamine synthetase, the proline rich complex, L-alanine dehydrogenase, phosphate binding protein, Ag 85 complex, HBHA (heparin binding hemagglutinin), MPT51, MPT64, superoxide dismutase, 19 kDa lipoprotein, α -crystallin, GroES, ESAT-6 when the first amino acid sequence is that of MPT59, and MPT59 when the first amino acid sequence is that of ESAT-6.
- 10 18. A fusion polypeptide fragment according to any of claims 11-17, which comprises ESAT-6 fused to MPT59.
19. A fusion polypeptide fragment according to claim 18, wherein no linkers are introduced between the two amino acid sequences.
- 15 20. A polypeptide according to any of the preceding claims which is lipidated so as to allow a self-adjuvating effect of the polypeptide.
21. A substantially pure polypeptide according to any of claims 1-20 for use as a pharmaceutical.
- 20 22. The use of a substantially pure polypeptide according to any of claims 1-20 in the preparation of a pharmaceutical composition for the diagnosis of or vaccination against tuberculosis caused by *Mycobacterium tuberculosis*, *Mycobacterium africanum* or *Mycobacterium bovis*.
- 25 23. A nucleic acid fragment in isolated form which
- 1) comprises a nucleic acid sequence which encodes a polypeptide as defined in any of claims 1-20, or comprises a nucleic acid sequence complementary thereto,
 - 2) has a length of at least 10 nucleotides and hybridizes
- 30 readily under stringent hybridisation conditions with a

nucleic acid fragment which has a nucleotide sequence selected from

SEQ ID NO: 1 or a sequence complementary thereto,
SEQ ID NO: 3 or a sequence complementary thereto,
5 SEQ ID NO: 5 or a sequence complementary thereto,
SEQ ID NO: 7 or a sequence complementary thereto,
SEQ ID NO: 9 or a sequence complementary thereto,
SEQ ID NO: 11 or a sequence complementary thereto,
SEQ ID NO: 13 or a sequence complementary thereto,
10 SEQ ID NO: 15 or a sequence complementary thereto,
SEQ ID NO: 41 or a sequence complementary thereto,
SEQ ID NO: 47 or a sequence complementary thereto,
SEQ ID NO: 49 or a sequence complementary thereto,
SEQ ID NO: 51 or a sequence complementary thereto,
15 SEQ ID NO: 53 or a sequence complementary thereto,
SEQ ID NO: 55 or a sequence complementary thereto,
SEQ ID NO: 57 or a sequence complementary thereto,
SEQ ID NO: 59 or a sequence complementary thereto,
SEQ ID NO: 61 or a sequence complementary thereto,
20 SEQ ID NO: 63 or a sequence complementary thereto,
SEQ ID NO: 65 or a sequence complementary thereto,
SEQ ID NO: 67 or a sequence complementary thereto,
SEQ ID NO: 69 or a sequence complementary thereto,
SEQ ID NO: 71 or a sequence complementary thereto,
25 SEQ ID NO: 87 or a sequence complementary thereto,
SEQ ID NO: 89 or a sequence complementary thereto,
SEQ ID NO: 91 or a sequence complementary thereto,
SEQ ID NO: 93 or a sequence complementary thereto,
SEQ ID NO: 140 or a sequence complementary thereto,
30 SEQ ID NO: 142 or a sequence complementary thereto,
SEQ ID NO: 144 or a sequence complementary thereto,
SEQ ID NO: 146 or a sequence complementary thereto,
SEQ ID NO: 148 or a sequence complementary thereto,
SEQ ID NO: 150 or a sequence complementary thereto, and
35 SEQ ID NO: 152 or a sequence complementary thereto,

with the proviso that when the nucleic acid fragment comprises a subsequence of SEQ ID NO: 41, then the nucleic acid

fragment contains an A corresponding to position 781 in SEQ ID NO: 41 and when the nucleic acid fragment comprises a subsequence of a nucleotide sequence exactly complementary to SEQ ID NO: 41, then the nucleic acid fragment comprises a T
5 corresponding to position 781 in SEQ ID NO: 41.

24. A nucleic acid fragment according to claim 23, which is a DNA fragment.

25. A vaccine comprising a nucleic acid fragment according to claim 23 or 24, the vaccine effecting in vivo expression of
10 antigen by an animal, including a human being, to whom the vaccine has been administered, the amount of expressed antigen being effective to confer substantially increased resistance to infections with mycobacteria of the tuberculo-
sis complex in an animal, including a human being.

15 26. A nucleic acid fragment according to claim 23 or 24 for use as a pharmaceutical.

27. The use of a nucleic acid fragment according to claim 23 or 24 in the preparation of a pharmaceutical composition for the diagnosis of or vaccination against tuberculosis caused
20 by *Mycobacterium tuberculosis*, *Mycobacterium africanum* or *Mycobacterium bovis*.

28. An immunologic composition comprising a polypeptide according to any of claims 1-20.

29. An immunologic composition according to claim 28, which
25 further comprises an immunologically and pharmaceutically acceptable carrier, vehicle or adjuvant.

30. An immunologic composition according to claim 29, wherein the carrier is selected from the group consisting of a polymer to which the polypeptide(s) is/are bound by hydrophobic
30 non-covalent interaction, such as a plastic, e.g. polystyrene, a polymer to which the polypeptide(s) is/are covalently

bound, such as a polysaccharide, and a polypeptide, e.g. bovine serum albumin, ovalbumin or keyhole limpet hemocyanin; the vehicle is selected from the group consisting of a diluent and a suspending agent; and the adjuvant is selected from the group consisting of dimethyldioctadecylammonium bromide (DDA), Quil A, poly I:C, Freund's incomplete adjuvant, IFN- γ , IL-2, IL-12, monophosphoryl lipid A (MPL), and muramyl dipeptide (MDP).

31. An immunologic composition according to any of claims 28 to 30, comprising at least two different polypeptide fragments, each different polypeptide fragment being a polypeptide according to any of claims 1-20.

32. An immunologic composition according to claim 31, comprising 3-20 different polypeptide fragments, each different polypeptide fragment being according to any of claims 1-20.

33. An immunologic composition according to any of claims 28-32, which is in the form of a vaccine.

34. An immunologic composition according to any of claims 28-32, which is in the form of a skin test reagent.

35. A vaccine for immunizing an animal, including a human being, against tuberculosis caused by mycobacteria belonging to the tuberculosis complex, comprising as the effective component a non-pathogenic microorganism, wherein at least one copy of a DNA fragment comprising a DNA sequence encoding a polypeptide according to any of claims 1-20 has been incorporated into the genome of the microorganism in a manner allowing the microorganism to express and optionally secrete the polypeptide.

36. A vaccine according to claim 35, wherein the microorganism is a bacterium.

37. A vaccine according to claim 36, wherein the bacterium is selected from the group consisting of the genera *Mycobacterium*, *Salmonella*, *Pseudomonas* and *Escherichia*.
38. A vaccine according to claim 37, wherein the microorganism is *Mycobacterium bovis* BCG, such as *Mycobacterium bovis* BCG strain: Danish 1331.
39. A vaccine according to any of claims 35-38, wherein at least 2 copies of a DNA fragment encoding a polypeptide according to any of claims 1-20 are incorporated into the genome of the microorganism.
40. A vaccine according to claim 39, wherein the number of copies is at least 5.
41. A replicable expression vector which comprises a nucleic acid fragment according to claim 23 or 24.
42. A vector according to claim 41, which is selected from the group consisting of a virus, a bacteriophage, a plasmid, a cosmid, and a microchromosome.
43. A transformed cell harbouring at least one vector according to claim 41 or 42.
44. A transformed cell according to claim 43, which is a bacterium belonging to the tuberculosis complex, such as a *M. tuberculosis bovis* BCG cell.
45. A transformed cell according to claim 43 or 44, which expresses a polypeptide according to any of claims 1-20.
46. A method for producing a polypeptide according to any of claims 1-20, comprising
- inserting a nucleic acid fragment according to claim 23 or 24 into a vector which is able to replicate in a host cell,

introducing the resulting recombinant vector into the host cell, culturing the host cell in a culture medium under conditions sufficient to effect expression of the polypeptide, and recovering the polypeptide from the host
5 cell or culture medium; or

isolating the polypeptide from a short-term culture filtrate as defined in claim 1; or

isolating the polypeptide from whole mycobacteria of the tuberculosis complex or from lysates or fractions thereof,
10 e.g. cell wall containing fractions; or

synthesizing the polypeptide by solid or liquid phase peptide synthesis.

47. A method for producing an immunologic composition according to any of claims 28-32 comprising

15 preparing, synthesizing or isolating a polypeptide according to any of claims 1-20, and

solubilizing or dispersing the polypeptide in a medium for a vaccine, and

20 optionally adding other *M. tuberculosis* antigens and/or a carrier, vehicle and/or adjuvant substance,

or

cultivating a cell according to any of claims 37-45, and

transferring the cells to a medium for a vaccine, and

25 optionally adding a carrier, vehicle and/or adjuvant substance.

48. A method of diagnosing tuberculosis caused by *Mycobacterium tuberculosis*, *Mycobacterium africanum* or *Mycobacterium bovis* in an animal, including a human being, comprising intradermally injecting, in the animal, a polypeptide according to any of claims 1-20 or an immunologic composition according to claim 34, a positive skin response at the location of injection being indicative of the animal having tuberculosis, and a negative skin response at the location of injection being indicative of the animal not having tuberculosis.

49. A method for immunising an animal, including a human being, against tuberculosis caused by mycobacteria belonging to the tuberculosis complex, comprising administering to the animal the polypeptide according to any of claims 1-20, the immunologic composition according to claim 33, or the vaccine according to any of claims 35-40.

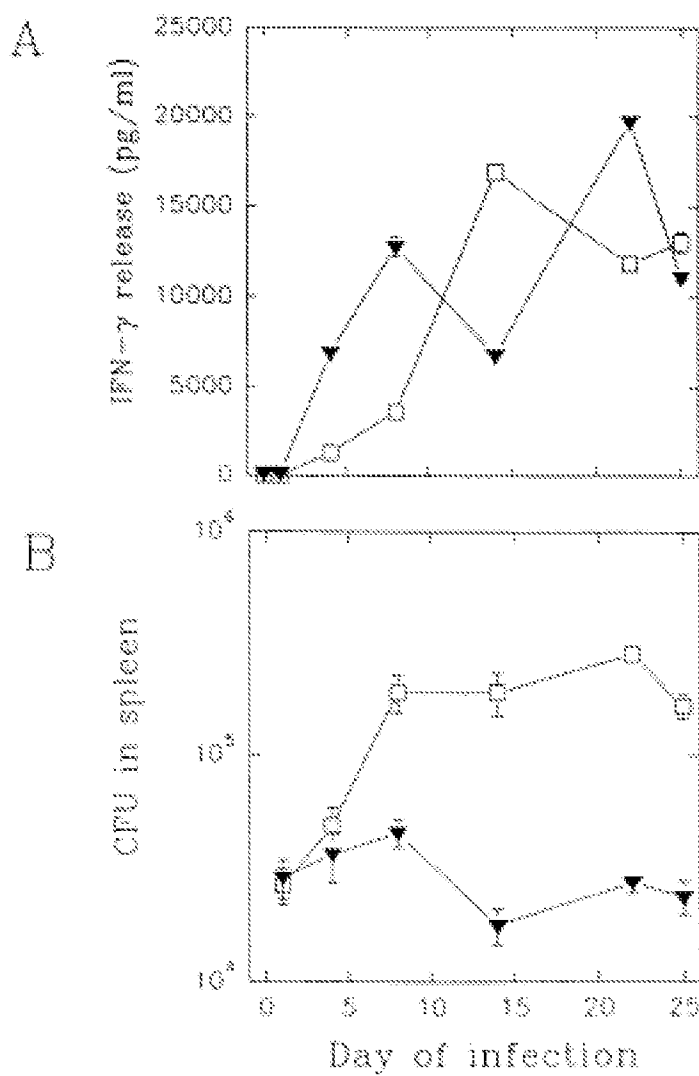
50. A method according to claim 49, wherein the polypeptide, immunologic composition, or vaccine is administered by the parenteral (such as intravenous and intraarterially), intraperitoneal, intramuscular, subcutaneous, intradermal, oral, buccal, sublingual, nasal, rectal or transdermal route.

51. A method for diagnosing ongoing or previous sensitization in an animal or a human being with bacteria belonging to the tuberculosis complex, the method comprising providing a blood sample from the animal or human being, and contacting the sample from the animal with the polypeptide according to any of claims 1-20, a significant release into the extracellular phase of at least one cytokine by mononuclear cells in the blood sample being indicative of the animal being sensitized.

52. A composition for diagnosing tuberculosis in an animal, including a human being, comprising a polypeptide according to any of claims 1-20, or a nucleic acid fragment according to claim 23 or 24, optionally in combination with a means for detection.

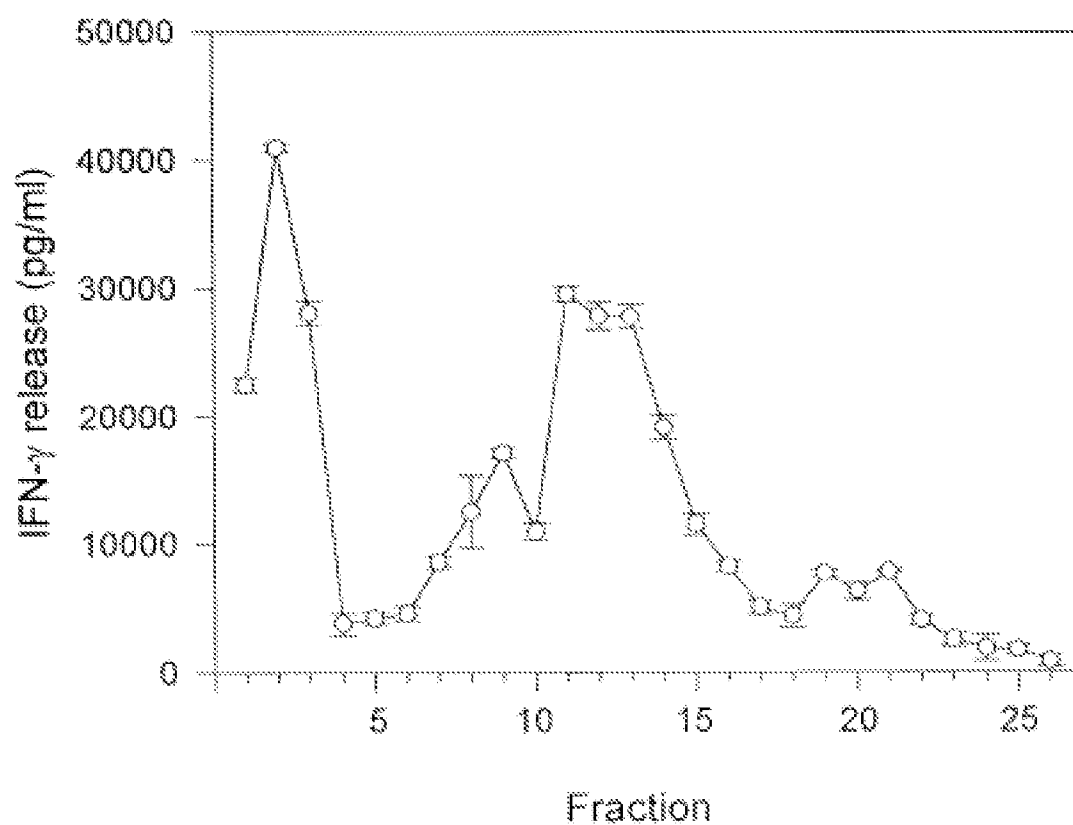
53. A monoclonal or polyclonal antibody, which is specifically reacting with a polypeptide according to any of claims 1-20 in an immuno assay, or a specific binding fragment of said antibody.

1/6

**Fig. 1**

SUBSTITUTE SHEET (RULE 26)

2/6

**Fig. 2**

3/6

1 GGCCTGGCCGGT ACCATATGTTG CCGCCGATGC TCCGNCGCG TCGACCTATA CCGCCTTCG
 -35 region -10 region
 60
 61 ATCGAACCT GCTGACCGAG AGGACTTGTG ATG TCG CAA ATC ATG TAC AAC TAC CCC GCG
 Shine Delgarno M S Q I M Y N Y P A
 120
 121 ATG TTG GGT CAC GCC GCG GAT ATG GCC GGA TAT GCC GGC ACC CTG CAG AGC TTG GGT GCC
 M L G H A G D M A G Y A G T L Q S L G A
 180
 181 GAG ATC GCC GTG CAG CAG GCC GCG TTG CAG AGT GCG TCG CAG GGC GAT ACC GGG ATC ACG
 E I A V E Q A A L Q S A W Q G D T G I T
 240
 241 TAT CAG GCG TCG CAG GCA CAG TCG AAC CAG GCC ATG GAA GAT TTG GTG CCG GCC TAT CAT
 Y Q A W Q A Q W N Q A M E D L V R Y H A
 300
 301 GCG ATG TCG AGC ACC CAT GAA GCC AAC ACC ATG GCG ATG ATG GCC CCG GAC ACC GCC GAA
 Y M S S T H E A N T M A M M A R D T A E
 360
 361 GCC GCC AAA TCG GGC GGC TAG
 A A X W G G *

Fig. 3

1 GGGTACCCCG ACCACGGCTG GGCAGAGATG TGCAGGCGGC CATCAAGGCG GTCAAGGCGG 60
 -35 region
 61 GCGACGGCGT CATATACCTG GACGGCACCT TGTTCGGCGG CCGCGGGTGT CTGACGCGCG 120
 -10 region
 121 ACCAGTACAA CTCGCCGCTG GTG GCC GCC GAC CCG GAG TCC ACC GCG GCG 170
 Shine Dalgarno V A A D P E S T A A
 171 TTG CCC GAC GGC GCC GGG CTG GTC GTT CTG GAT GGC ACC GTC ACT GCC GAA CTC GAA GCC 230
 L P D G A G L V V L D G T V T A E L E A
 231 GAG GGC TGG GCC AAA GAT CGC ATC CGC GAA CTG CAA GAG CTG CGT AAG TCG ACC GCG CTG 290
 E G W A K D R I R E L Q E L R K S T G L
 291 GAC GTT TCC GAC CGC ATC CGG GTG GTG ATG TCG GTG CCT CGG GAA CGC GAA GAC TGG GCG 350
 D V S D R I R V V M S V P A E R E D W A
 351 CGC ACC CAT CGC GAC CTC ATT GCC GGA GAA ATC TTG GCT ACC GAC TTC GAA TTC GCC GAC 410
 R T H R D L I A G E I L A T D F E F A D
 411 CTC GCC GAT GGT GTG GCC ATC GGC GAC GGC GTG CGG GTA AGC ATC GAA AAG ACC TGA 467
 L A D G V A I G D G V R V S I E K T *

Fig. 4

5/6

1	GAATTCGCCCGGTGCACACAGCCTTACACGACGGAGGTTGGACACATGAAG	50
	M K	
51	GGTCGGTCGGCGCTGCTGCGGGCGCTCTGGATTGCCGCACTGTCATTCCG	100
	G R S A L L R A L W I A A L S F G	
101	GTTGGGCGGTGTGCGCGGTAGCCGCGGAACCCACCGCCAAGGCCGCCCAT	150
	L G G V A V A A E P T A K A A P	
151	ACGAGAACCTGATGGTGCCGTCGCCCTCGATGGGCCGGGACATCCCGGTG	200
	Y E N L M V P S P S M G R D I P V	
201	GCCTTCCTAGCCGGTGGGCGCGCACCGCGGTGTATCTGCTGGACGCCITCAA	250
	A F L A G G P H A V Y L L D A F N	
251	CGCCGGCCCCGGATGTCAGTAACTGGGTACCGCGGGTAACGCGGATGAACA	300
	A G P D V S N W V T A G N A M N	
301	CGTGGCGGGCAAGGGGATTTCCGGTGGTGGCACCGGCCGGTGGTGCGTAC	350
	T L A G K G I S V V A P A G G A Y	
351	AGCATGTACACCAACTGGGAGCAGGATGGCAGCAAGCAGTGGGACACCTT	400
	S M Y T N W E Q D G S K Q W D T F	
401	CTTGTCGCTGAGCTGCCCGACTGGCTGGCCGCTAACCGGGGCTTGGCCC	450
	L S A E L P D W L A A N R G L A	
451	CCGGTGGCCATGCCGGCCGTTGGCGCCCGCTCAGGGCGCGTTACGGGGCGGATG	500
	P G G H A A V G A A Q G G Y G A M	
501	GCGCTGGCGGCCTTCCACCCCGACCGCTTCGGCTTCGCTGGCTCGATGTC	550
	A L A A F H P D R F G F A G S M S	
551	GGGCTTTTTGTACCCGTCGAACACCACCACCAACGGTGGGATCGCGGCGG	600
	G F L Y P S N T T T N G A I A A	
601	GCATGCAGCAATTCCGGCGGTGTGGACACCAACGGAATGTGGGGAGCACCA	650
	G M Q Q F G G V D T N G M W G A P	
651	CAGCTGGGTGCGGTGGAAGTGGCAGCACCGCTGGGTGCATGCCAGCCTGCT	700
	Q L G R W K W H D P W V H A S L L	
701	GGCGCAAAACAACACCCGGGTGTGGGTGTGGAGCCCCGACCAACCCGGGAG	750
	A Q N N T R V W V W S P T N P G	
751	CCAGCGATCCCCGCCCGCCATGATCGGCCAAACCGCCGAGGCGATGGGTAAC	800
	A S D P A A M I G Q T A E A M G N	
801	AGCCGCATGTTCTACAACCAGTATCGCAGCGTCCGGCGGGCACAACGGACA	850
	S R M F Y N Q Y R S V G G H N G H	
851	CTTCGACTTCCCAGCCAGCGGTGACAACGGCTGGGGCTCGTGGGGCGCCCC	900
	F D F P A S G D N G W G S W A P	
901	AGCTGGGCGCTATGTGCGGGCGATATCGTCGGTGGGATCCGCTAAGCGAAT	950
	Q L G A M S G D I V G A I R	
951	TC	952

Fig. 5

6/6

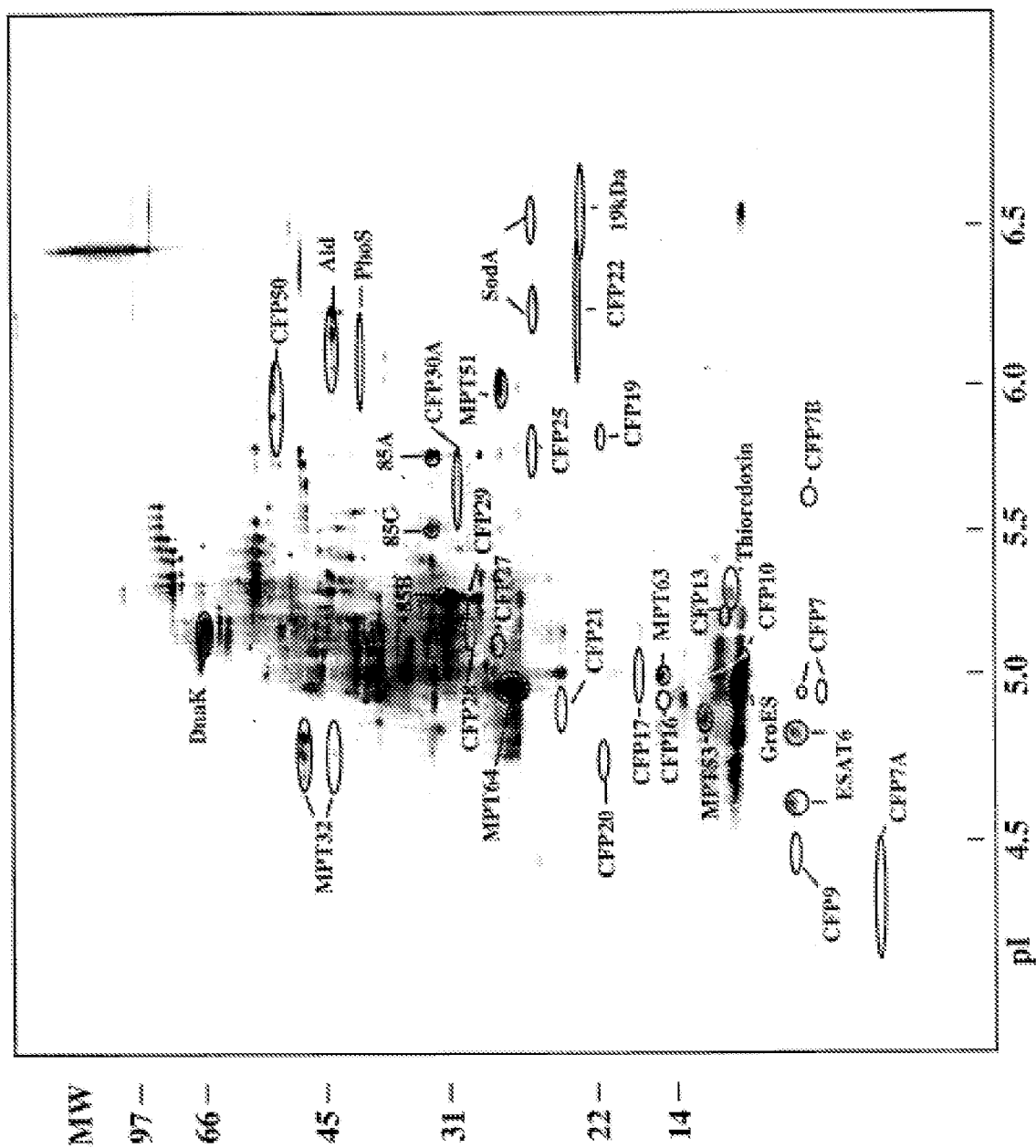


Fig. 6

SUBSTITUTE SHEET (RULE 26)

INTERNATIONAL SEARCH REPORT

International Application No.

PCT/DK 98/00132

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/31 A61K39/04 C07K14/35 C12N15/62 A61K38/16
G01N33/569 C12Q1/68 C07K16/12

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N A61K C07K G01N C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
----------	--	-----------------------

X	<p>WD 97 09428 A (CORIXA CORP) 13 March 1997</p> <p>see the claims see abstract: examples 1,3 see page 12 - page 18 see page 21 - page 25</p> <p style="text-align: center;">--- -/-</p>	<p>1-4,6,7, 9-13,15, 16, 21-43, 45-53</p>
---	--	---

☒ Further documents are listed in the continuation of box C.

☒ Related family members are listed in annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another claim or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"S" document member of the same patent family

Date of the actual completion of the international search

1 September 1998

Date of mailing of the international search report

15. 09. 1999

Name and mailing address of the ISA

European Patent Office, P.O. Box 5018 Patentplan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 851 494 nl
Fax: (+31-70) 340-3018

Authorized officer

Oderwald, H

INTERNATIONAL SEARCH REPORT

Inventor's Application No.

PCT/DK 98/00132

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>WO 97 09429 A (CORIXA CORP) 13 March 1997</p> <p>see the claims see abstract; examples 1,3 see page 12 - page 15, paragraph 2 see page 17, paragraph 4 - page 19, paragraph 2 see page 24 - page 25</p>	<p>1-4,6,7, 9-13,15, 16, 21-43, 45-47, 49-53</p>
X	<p>WO 95 01441 A (STATENS SERUMSINSTITUT ;ANDERSEN PETER (DK); ANDERSEN AASE BENGGAAR) 12 January 1995</p> <p>see the claims see abstract; figure 10; examples 1,3-6; table 2 see page 12 - page 32 see page 12, paragraph 3</p>	<p>1-4,6,7, 20-53</p>
X	<p>SORENSEN A L ET AL: "Purification and characterization of a low-molecular-mass T-cell antigen secreted by Mycobacterium tuberculosis." INFECTION AND IMMUNITY, (1995 MAY) 63 (5) 1710-7. JOURNAL CODE: 607. ISSN: 0019-9567., XP002068818 cited in the application see abstract; figures 4-6 see page 1710, paragraph 3 - page 1712, paragraph 4 see page 1713, paragraph 5 see page 1716, paragraph 5 - paragraph 8</p>	<p>1-4,6,7, 9,10, 21-24, 28,33, 34, 41-43, 45-53</p>
X	<p>CRABTREE J AND ROE B A: "Homo sapiens clone 137c7" EMBL SEQUENCE DATABASE, 19 March 1997, XP002068854 HEIDELBERG, GERMANY see the whole document</p>	<p>23-27, 41-43</p>
X	<p>VALDES-STAUBER N AND SCHERER S: "Nucleotide sequence and taxonomical distribution of the bacteriocin gene lin cloned from Brevibacterium linens M18" APPLIED AND ENVIRONMENTAL MICROBIOLOGY, vol. 62, no. 4, April 1996, pages 1283-1286, XP002076056 see the whole document</p>	<p>1-4,6,8, 23,24, 41-43, 45,46,52</p>

-/-

INTERNATIONAL SEARCH REPORT

International Application No.

PCT/DK 98/00132

C. (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 96 37219 A (UNIV CALIFORNIA ; HORWITZ MARCUS A (US); HARTH GUENTER (US)) 28 November 1996 see the claims see abstract; figures 2-4, 13; examples 4, 6, 14, 21-26	
P, X	BROWN D AND CHURCHER C M: "Mycobacterium tuberculosis cosmid v035" EMBL SEQUENCE DATABASE, 20 February 1998, XP002068855 HEIDELBERG, GERMANY see the whole document	23-27, 41-43
T	ROSENKRANDS I ET AL: "Identification and characterization of a 29-kilodalton protein from Mycobacterium tuberculosis culture filtrate recognized by mouse memory effector cells" INFECTION AND IMMUNITY, vol. 66, no. 6, June 1998, pages 2718-2735, XP002076057 see abstract; figure 4 see page 2728, paragraph 4 - page 2729, paragraph 20	1-4, 6-9, 23, 24, 41-46, 52, 53
P, X	& ROSENKRANDS I ET AL: "CFP29 protein (accession number D07812)" EMBL SEQUENCE DATABASE, 1 July 1997, Heidelberg, Germany see the whole document	1-4, 6-8
P, X	& ROSENKRANDS I ET AL: "Mycobacterium tuberculosis cfp29 gene (accession number Y12820)" EMBL SEQUENCE DATABASE, 30 June 1997, Heidelberg, Germany see the whole document	23, 24, 41-46, 52

INTERNATIONAL SEARCH REPORT

International application No.
PCT/OK 98/00132

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
Although claims 49 and 50 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.
2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☒ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
1-4,6-17,20-53; inventions 1 and 8
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
☒ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-4, 6-17, 20-53 all partially

A polypeptide fragment from mycobacteria belonging to the tuberculosis complex comprising the amino acid SEQ ID NO: 2, nucleic acids encoding said polypeptide as in SEQ ID NO:1, fusion proteins comprising said polypeptides, vaccines, pharmaceutical and immunological compositions containing said polypeptide or nucleic acid, an expression vector comprising said nucleic acid, a host transformed with said vector, immunization with said polypeptide, the use of said polypeptide in diagnosis, antibodies against said polypeptide.

2. Claims: 1-4, 6-17, 20-53 all partially

same as invention 1 but for SEQ ID NO: 4 and 3.

3. Claims: 1-17, 20-53 all partially

same as invention 1 but for SEQ ID NO: 6, 5 and 17.

4. Claims: 1-4, 6-17, 20-53 all partially

same as invention 1 but for SEQ ID NO: 8, 7 and 18.

5. Claims: 1-17, 20-53 all partially

same as invention 1 but for SEQ ID NO: 10, 9 and 19.

6. Claims: 1-17, 20-53 all partially

same as invention 1 but for SEQ ID NO: 12, 11 and 20.

7. Claims: 1-17, 20-53 all partially

same as invention 1 but for SEQ ID NO: 14, 13 and 21.

8. Claims: 1-4, 6-17, 20-53 all partially

same as invention 1 but for SEQ ID NO: 16, 15 and 23.

9. Claims: 1-4, 6-17, 20-53 all partially

same as invention 1 but for SEQ ID NO: 22.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

10. Claims: 1-17, 20-53 all partially
same as invention 1 but for SEQ ID NO: 42 and 41.
11. Claims: 1-4, 6-17, 20-53 all partially
same as invention 1 but for SEQ ID NO: 48, 47 and 81.
12. Claims: 1-4, 6-17, 20-53 all partially
same as invention 1 but for SEQ ID NO: 50, 49 and 82.
13. Claims: 1-17, 20-53 all partially
same as invention 1 but for SEQ ID NO: 52 and 51.
14. Claims: 1-4, 6-17, 20-53 all partially
same as invention 1 but for SEQ ID NO: 54, 53 and 83.
15. Claims: 1-17, 20-53 all partially
same as invention 1 but for SEQ ID NO: 56 and 55.
16. Claims: 1-17, 20-53 all partially
same as invention 1 but for SEQ ID NO: 58, 57 and 84.
17. Claims: 1-4, 6-17, 20-53 all partially
same as invention 1 but for SEQ ID NO: 60, 59 and 85.
18. Claims: 1-4, 6-17, 20-53 all partially
same as invention 1 but for SEQ ID NO: 62, 61 and 86.
19. Claims: 1-4, 6-17, 20-53 all partially
same as invention 1 but for SEQ ID NO: 64, 63 and 79.
20. Claims: 1-4, 6-17, 20-53 all partially
same as invention 1 but for SEQ ID NO: 66, 65 and 78.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

21. Claims: 1-4, 6-17, 20-53 all partially
same as invention 1 but for SEQ ID NO: 68 and 67.
22. Claims: 1-4, 6-17, 20-53 all partially
same as invention 1 but for SEQ ID NO: 70 and 69.
23. Claims: 1-4, 6-17, 20-53 all partially
same as invention 1 but for SEQ ID NO: 72 and 71.
24. Claims: 1-4, 6-17, 20-53 all partially
same as invention 1 but for SEQ ID NO: 75.
25. Claims: 1-4, 6-17, 20-53 all partially
same as invention 1 but for SEQ ID NO: 76.
26. Claims: 1-4, 6-17, 20-53 all partially
same as invention 1 but for SEQ ID NO: 80.
27. Claims: 1-4, 6-17, 20-53 all partially
same as invention 1 but for SEQ ID NO: 88 and 87.
28. Claims: 1-4, 6-17, 20-53 all partially
same as invention 1 but for SEQ ID NO: 90 and 89.
29. Claims: 1-4, 6-17, 20-53 all partially
same as invention 1 but for SEQ ID NO: 92 and 91.
30. Claims: 1-4, 6-17, 20-53 all partially
same as invention 1 but for SEQ ID NO: 94 and 93.
31. Claims: 1-4, 6-17, 20-53 all partially
same as invention 1 but for SEQ ID NO: 141, 140 and 169.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

32. Claims: 1-4, 6-17, 20-53 all partially
same in invention 1 but for SEQ ID NO: 143, 142 and 170.

33. Claims: 1-4, 6-17, 20-53 all partially
same as invention 1 but for SEQ ID NO: 145, 144 and 171.

34. Claims: 1-4, 6-17, 20-53 all partially
same as invention 1 but for SEQ ID NO: 147, 146 and 168.

35. Claims: 1-4, 6-17, 20-53 all partially
same as invention 1 but for SEQ ID NO: 149, 148 and 73.

36. Claims: 1-17, 20-53 all partially
same as invention 1 but for SEQ ID NO: 151, 150 and 74.

37. Claims: 1-4, 6-17, 20-53 all partially
same as invention 1 but for SEQ ID NO: 153, 152 and 77.

38. Claims: 11-17, 20-53 all partially, 18, 19

A fusion polypeptide comprising ESAT-6 or MPT59 each individually with one of the following epitope partners: DnaK, GroEL, urease, glutamine synthetase, the proline rich complex, L-alanine dehydrogenase, phosphate binding protein, Ag 85 complex, HBHA, MPT51, MPT64, superoxide dismutase 19 kDa lipoprotein, alpha-crystallin, GroES, nucleic acids encoding said polypeptide, vaccines, pharmaceutical and immunological compositions containing said polypeptide or nucleic acid, an expression vector comprising said nucleic acid, a host transformed with said vector, immunization with said polypeptide, the use of said polypeptide in diagnosis, antibodies against said polypeptide.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No.

PCT/DK 98/00132

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9709428 A	13-03-1997	AU 7158696 A EP 0851927 A NO 980883 A	27-03-1997 08-07-1998 27-04-1998
WO 9709429 A	13-03-1997	AU 7158796 A EP 0850305 A	27-03-1997 01-07-1998
WO 9501441 A	12-01-1995	AU 682879 B AU 7068894 A CA 2165949 A EP 0706571 A NZ 267984 A	23-10-1997 24-01-1995 12-01-1995 17-04-1996 22-09-1997
WO 9637219 A	28-11-1996	AU 6024596 A EP 0828510 A	11-12-1996 18-03-1998